Second Annual Report for the Land Information System

Submitted under Task Agreement GSFC-CT-2

Cooperative Agreement Notice (CAN) CAN-00OES-01

Increasing Interoperability and Performance of Grand Challenge Applications in the Earth, Space, Life, and Microgravity Sciences

Period Covered: September 1, 2002-August 31, 2003

Version 2.0

Table of Contents

List of Figures

List of Tables

Acronyms and Terms

ALMA: Assistance for Land-surface Modeling Activities

API: Application Programming Interface

CGI: Common Gateway Interface

CLM: Community Land Model

DODS: Distributed Ocean Data System

EMC: Environmental Modeling Center

ESMF: Earth System Modeling Framework

GFS: Global Forecast System

GrADS: Grid Analysis and Display System

LDAS: Land Data Assimilation System

LIS: Land Information System

MRTG: Multi Router Traffic Grapher

NCEP: National Centers for Environmental Prediction

NFS: Network File System

NOAH: National Centers for Environmental Prediction, Oregon State University, United States Air Force, and Office of Hydrology Land Surface Model

NOAA: National Oceanic and Atmospheric Administration

PXE: Preboot Execution Environment

RAID: Redundant Array of Inexpensive Disks

SNMP: Simple Network Management Protocol

VIC: Variable Infiltration Capacity Land Surface Model

1 Objective

This Annual Report summarizes the activities and accomplishments of the Land Information System (LIS) project during FY03. The overarching objective of the LIS project is to build a high-resolution, high-performance land surface modeling and data assimilation system to support a wide range of land surface research activities and applications.

This document has been prepared in accordance with the requirements of the Task Agreement GSFC-CT-2 under Cooperative Agreement Notice CAN-00-OES-01 Increasing Interoperability and Performance of Grand Challenge Applications in the Earth, Space, Life, and Microgravity Sciences, funded by NASA's ESTO Computational Technologies (formerly High Performance Computing and Communications) Project.

1.1 Project goals

As stated in the proposal and approved software engineering documents, the goals of LIS include:

- Realistic land surface modeling. LIS will simulate the global land surface variables using various land surface models, driven by atmospheric "forcing data" (e.g., precipitation, radiation, wind speed, temperature, humidity) from various sources.
- High performance computing. LIS will perform high-performance, parallel computing for near real-time, high-resolution land surface modeling research and operations.
- Efficient data management. The high-resolution land surface simulation will produce a huge data throughput, and LIS will retrieve, store, interpolate, re-project, sub-set, and backup the input and output data efficiently.
- Usability. LIS will provide intuitive web-based interfaces to users with varying levels of access to LIS data and system resources, and enforce user security policies.
- Interoperable and portable computing. LIS will incorporate the ALMA (Assistance for Land surface Modeling Activities) and ESMF (Earth System Modeling Framework) standards to facilitate inter-operation with other Earth system models. In order to demonstrate portability of LIS, the land surface modeling component will be implemented on a custom-designed Linux cluster and an SGI Origin 3000.

1.2 Land Surface Modeling Background

In general, land surface modeling seeks to predict the terrestrial water, energy and biogeochemical processes by solving the governing equations of the soil-vegetation-snowpack medium. Land surface data assimilation seeks to synthesize data and land surface models to improve our ability to predict and understand these processes. The ability to predict terrestrial water, energy and biogeochemical processes is critical for applications in weather and climate prediction, agricultural forecasting, water resources management, hazard mitigation and mobility assessment.

In order to predict water, energy and biogeochemical processes using (typically 1-D vertical) partial differential equations, land surface models require three types of inputs: 1) initial conditions, which describe the initial state of land surface; 2) boundary conditions, which describe both the upper (atmospheric) fluxes or states—also known as "forcings"—and the lower (soil) fluxes or states; and 3) parameters, which are a function of soil, vegetation, topography, etc., and are used to solve the governing equations.

The LIS now includes various components that facilitate global land surface modeling within a data assimilation system framework. The main software components of the system are:

- Land Data Assimilation System (LDAS): LDAS is a software system that integrates land surface models in a data assimilation framework.
- Land surface Models: LIS includes 3 different land surface models, namely, CLM, NOAH, and VIC.

These components are explained in detail in the LIS Software Design Document.

2 Approach

The overall approach for the LIS project involves parallel efforts throughout the three-year project directed towards the performance and interoperability goals of the original CAN, while adopting standard software engineering methodologies to ensure that the products are portable, reusable, extensible and maintainable.

Activities during FY03 have been focused on progress toward the interoperability and first code improvement milestones. An annual "all investigator" meeting, including interoperability and user interface design reviews, was held on January 21, 2003, in addition to continuing weekly local team meetings and monthly teleconferences. We also recently became formal "ESMF partners", and have actively participated in ESMF meetings and reviews. The FY03 milestones and out-year milestones are given in the table below, and are described in more detail on the LIS web site (http://lis.gsfc.nasa.gov).

2.1 FY03 Milestones

H) Design policy for Interoperability and Community Delivery	28-FEB-2003
F) First Code Improvement	30-MAR-2003
I) Interoperability Prototype	30-JUL-2003
C) FY03 Annual Report	30-AUG-2003

2.2 Out-year Milestones

G) Second Code Improvement	28-FEB-2004
J) Full Interoperability demonstrated	30-JUL-2004
K) Customer delivery Accomplished	30-AUG-2004
K) Present LIS to Review Board	30-AUG-2004
D) Final Report Delivered	28-FEB-2005

3 Scientific Accomplishments

The Land Information System is based largely on the Land Data Assimilation System (LDAS; http://ldas.gsfc.nasa.gov), and is the product of a 5-year effort towards the goal of modeling land surface states and fluxes, while relying as much as possible on observation-based parameter and forcing fields in order to avoid biases that are known to exist in forcing fields produced by atmospheric models. A recently published article describes the science of LDAS in detail (Rodell et al., 2003).

4 Technology Accomplishments

As stated above, the second year of the LIS project has focused on progress toward the FY03 interoperability and code improvement milestones, and the technology accomplishments associated with these milestones are described below.

4.1 Design Policy for Interoperability and Community Delivery

As stated above, the LIS interoperability design was reviewed by members of the LIS and LDAS teams, as well as collaborators from Princeton, COLA, and NCEP, at the LIS annual meeting held January 21, 2003. The LIS interoperability design was approved by CT in March 2003. As part of the Milestone H submission, the following documents were submitted and approved:

- Software Design Document, including detailed “sub-documents”:
    - User Interface Design Document
    - Interface Design for Interoperability
    - Data Management Document
- Software Requirements Specification
- Software Traceability Matrix

As described in detail in the above documents, LIS defines two different types of interoperability: internal and external. ***Internal*** interoperability mainly deals with the aspects of making components within LIS interoperable, and ***external*** interoperability deals with the interaction of LIS with other related scientific community applications and standards. Internal interoperability within LIS is achieved by redesigning the LDAS driver to provide flexible and adaptive interfaces between subsystems, including using advanced features of the Fortran 90 programming language to simulate object oriented behavior and thereby provide well-defined interfaces or "hook points" for adding additional land surface models and input data sets. External interoperability is achieved by adopting the ALMA data exchange convention and by defining LIS as a fully ESMF-compliant land model component.

4.2 First Code Improvement

The first code improvement milestone for LIS was approved in April, 2003, and included the first application of the global LIS at a 5km spatial resolution. The milestone was achieved on the NASA Ames SGI Origin system "Lomax", based on the heavy use of Chapman for the Columbia accident investigation and the creation of a special queue on Lomax for the LIS team. The Milestone F results demonstrate significant improvement in the code relative to the Milestone E baseline, including parallelization of the previously serial baseline code and P/2 scaling up to 8 processors, as shown in Figure 1.

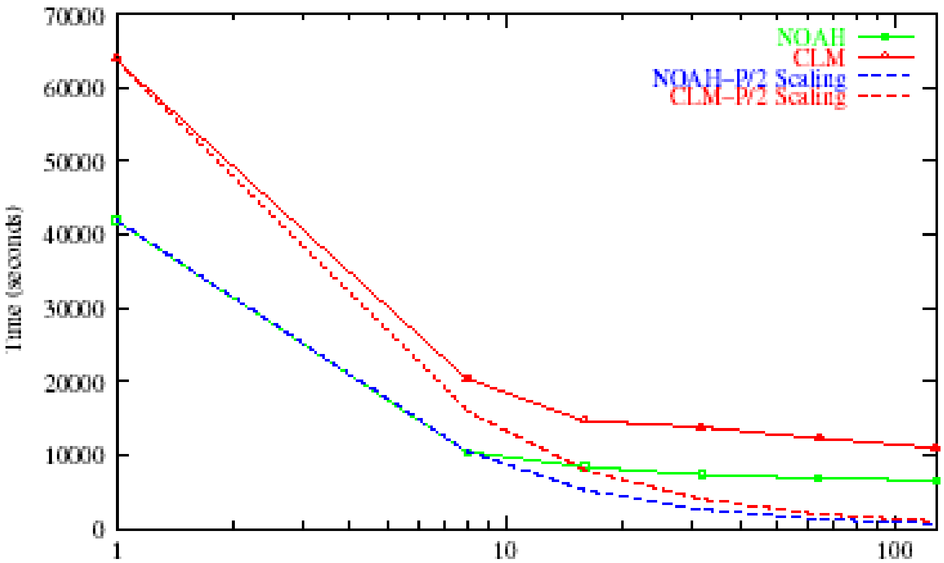

Figure 1. LIS timing results for Milestone F on Lomax at 5km.

Based on our Milestone E baselining results, the code improvements for Milestone F concentrated on improving the performance of functions for spatial and temporal interpolation, as well as memory management within the land surface models themselves. Optimization of these routines, including a substantial memory reduction due to the removal of ocean points from the LIS code, resulted in a factor of 2 speedup in runtime at ¼ degree spatial resolution for a serial execution on the Ames SGI Origin 3000 system Lomax, as shown in Table 1.

Table 1. Reduction in execution time on a single processor (NAS SGI O3K "Lomax") for LIS v1.0 relative to LDAS Baseline for one day, 1/4 degree global run with the Noah land surface model.

	Wallclock time
	(minutes)
LDAS Baseline (Milestone E)	63.4
LIS v1.0 (Milestone F)	27.4
reduction factor	**2.3**

Further testing within the LIS/LDAS teams was conducted to quantify the performance of the LIS code base relative LDAS code base under parallel development in the Hydrological Sciences Branch. As part of this testing, the May 2003 version of LDAS and the LIS v1.0 codes were executed in serial mode (as in Table 1) on a dedicated SGI Origin200 system "dew", within the Hydrological Sciences Branch at NASA/GSFC. Table 2 illustrates even greater performance gains of the LIS v1.0 code base relative to the May 2003 LDAS code, including a factor of 5 speedup in runtime, and a factor of 10 reduction in memory.

Table 2. Reduction in memory and execution time on a single processor (GSFC SGI O200 "dew") for LIS v1.0 relative to LDAS May 2003 version for one day, 1/4 degree global run with the Noah land surface model.

	Memory (MB)	Wallclock time (minutes)	CPU time (minutes)
LDAS May 2003 Version	3169	116.7	115.8
LIS v1.0 (Milestone F)	313	22.0	21.8
reduction factor	**10.1**	**5.3**	**5.3**

4.3 Interoperability Prototype

The LIS Interoperability Prototype for Milestone I was submitted to the CT team on August 15, 2003. Based on feedback received on September 12, 2003, the associated software engineering documents are being revised for resubmission and approval. The LIS interoperability prototype demonstrates both internal and external interoperability as defined above, in addition to portability of the LIS code to the LIS cluster. In particular, there are three key tests demonstrated by the prototype:

1. Internal Interoperability Test

   Summary: Demonstrate interoperability of LIS driver by adding the third Land Surface Model, VIC. Demonstrate that LIS/VIC combination works on SGI 3000.
2. External Interoperability Test

   Summary: Verify that ALMA mandatory outputs are available from CLM, Noah, and VIC in the current LIS driver, including variable names, units and sign conventions.
3. Portability Test

   Summary: Verify that LIS will run on the LIS Cluster at both ¼ degree and 5 km resolution. In this test, only the ¼ degree runs were successful due to a bug found in the GDS code.

As shown in Figures 2 and 3, the 5km VIC results on the SGI suggest significant new capabilities from the LIS, particularly in the area of resolving the effects of urban areas on energy fluxes, commonly known as the "urban heat island".

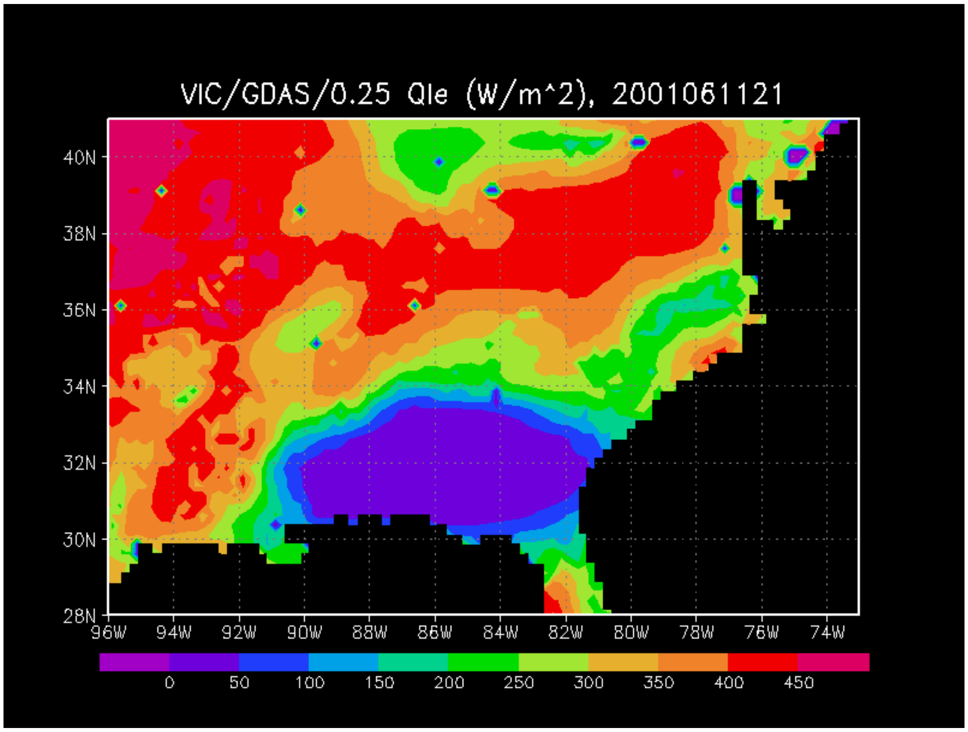

Figure 2. Latent heat flux (Qle) from the VIC land surface model in LIS v2.0. Spatial resolution is 1/4 degree.

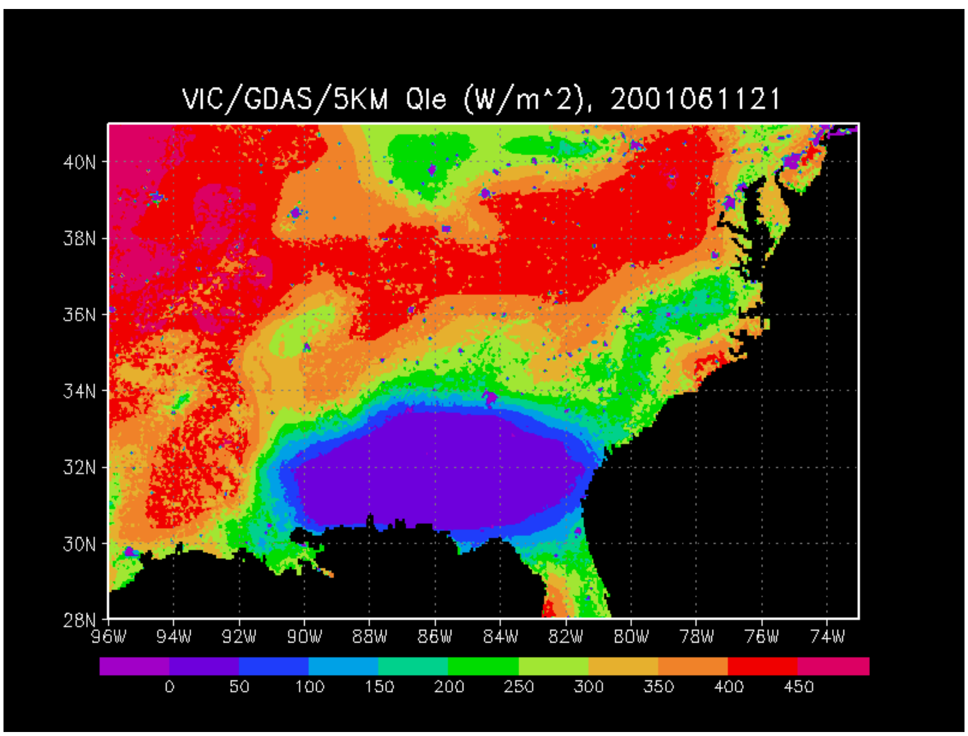

Figure 3. Same as Figure 2., except spatial resolution is 5 km. Note features resolved by the increased resolution, including urban areas (e.g., Baltimore-Washington, Atlanta, Houston), as well as the coastlines.

5 Status/Plans

In addition to the Technological Accomplishments described above, the LIS team has made significant progress towards future milestones in the areas of interfacing with ESMF, user interface development and customer delivery. Key progress in these areas is described below, followed by a discussion of plans for FY04.

5.1 Interfacing with the developing ESMF

In an effort to ensure that the developing ESMF meets the needs of land surface modeling, and will provide capabilities that can be successfully demonstrated in LIS, we have formed several linkages with the ESMF team, including becoming ESMF "Partners", with access to interim releases of the ESMF code. In addition, we have attended all ESMF team meetings, and have fully participated in component reviews, including the most recent Time Manager Interface Review. Further, Arlindo da Silva continues to be a member of our review board, and Paul Houser is a member of the ESMF review board.

The most troubling aspect of the LIS team's dependency on ESMF is the expectation that the ESMF team's release of an ESMF-compliant version of CLM will be delayed until at least Spring, 2004, which puts our July 2004 Milestone J in jeopardy. To mitigate this risk, the LIS team proposes that our interoperability design, which includes the definition of LIS itself (including CLM, VIC and NOAH) as an ESMF land surface model component, will meet the spirit of the full interoperability milestone without relying on the ESMF team's ability to meet their own demonstration milestone in a timely manner. The LIS team would like to propose a teleconference to further discuss this issue, and renegotiate our milestones as appropriate and necessary.

5.2 User Interface Development

As described in our Software Design document, and the associated User Interface Design document, the LIS user interface will provide functionality for three levels of users, including associated data access and security requirements. Level 1 users are the general public, who will access the LIS data through a standard web browser, as shown in Figure 4. Information provided to this class will only include static images, animations and text.

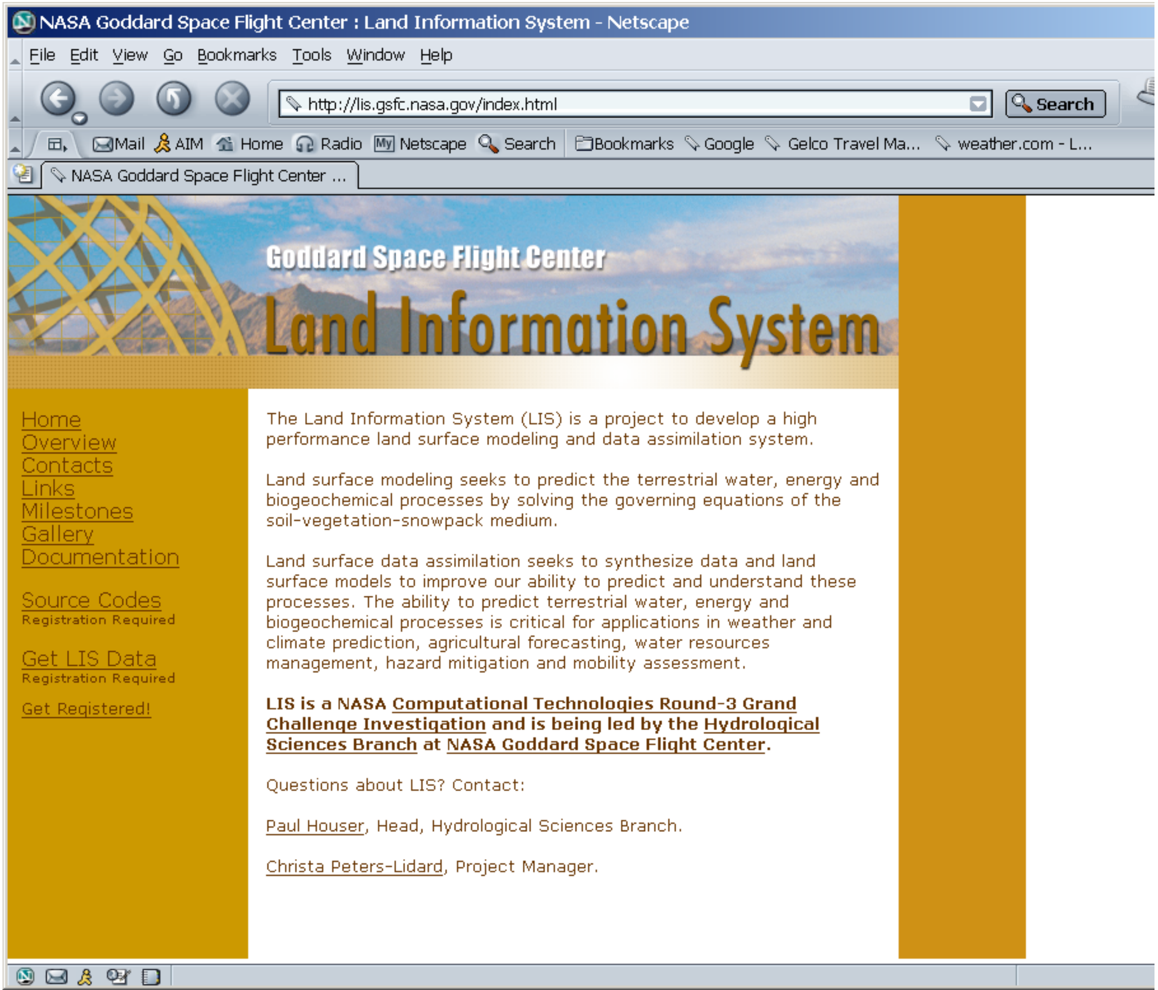

Figure 4. The Land Information System Level 1 User Interface.

Level 2 users will have more direct access to LIS data, and therefore must register using the links provided in the Level 1 interface shown in Figure 4. Level 2 users will be allowed to either dynamically generate an image via a web browser; use a DODS client; or access raw data from the LIS ftp server.

Level 3 users have the highest access level in addition to all the Level 2 and Level 1 access privileges; they will be able to access the parallel computing power of LIS system, including an account on the LIS cluster and a web interface for submitting LIS jobs, as shown in Figure 5. Level 3 users will be the LIS developers and collaborators, and a select group of users.

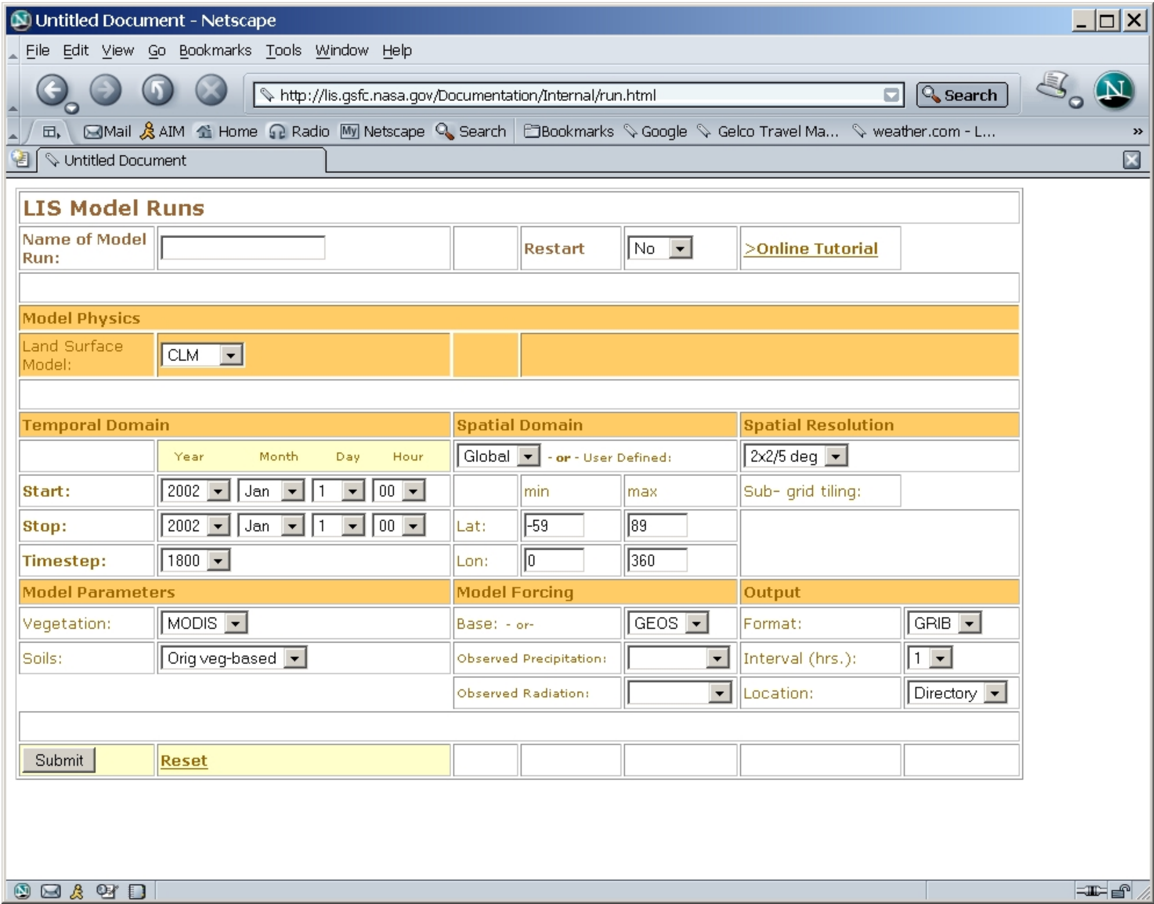

Figure 5. Prototype LIS level 3 user interface.

5.3 Customer Delivery to NOAA/NCEP

Towards our August 2004 customer delivery Milestone K, the LIS team has initiated an effort to port LIS v2.0 (Milestone I) to NOAA/NCEP for testing as an operational land data assimilation system. On May 29, 2003, the respective Land Teams of NCEP/EMC (Lead, Dr. Kenneth Mitchell) and NASA Goddard Hydrological Sciences Branch held a 1-day workshop at Goddard to begin the process of transferring LIS to the NCEP Central Computing Facility (NCCF, an IBM-SP2). The purpose is for applications by the EMC Land Team for initializing the land states of the NCEP/EMC Global Forecast System (GFS: both weather and seasonal) via various land data assimilation scenarios.

The two land teams have decided to carry out our first NCEP tests of LIS using the NCEP Noah land model for the test period of May 2003, on the GFS T62 Guassian grid, corresponding to the T62 Gaussian grid of EMC's current seasonal forecast model. In these tests LIS will use the same terrain-height field and land mask as the NCEP T62 global model. Moreover, these NCEP LIS tests will use exactly the same land surface characteristics fields (vegetation, soils) and parameters and Noah version number as NCEP's current tests of the Noah land model in the GFS. We fully expect that these efforts will contribute to LIS's portability and customer delivery objectives.

5.5 FY04 Plans

As shown above, the key FY04 milestones relate to performance, interoperability and customer delivery. Towards these milestones, we will be continuing to re-design and implement parallel I/O, based on the GDS, and in consultation with the CT staff scientists, in order to achieve the second code improvement milestone. In addition, we will fully implement our Level 2 and Level 3 user interfaces, including web configuration and visualization interfaces, GDS-LDAS interface, and LDAS-land surface model interfaces. We will also continue the porting and testing of the LIS code to NCEP for eventual use as an operational system. Finally, further adoption of the ESMF infrastructure utilities will continue, based on interim releases, in order to support the definition of LIS as a fully-compliant land model component.

6 Points of Contact

Dr. Paul Houser, Co-PI
Head, Hydrological Sciences Branch
NASA-GSFC Code 974
Greenbelt, Maryland 20771
Office: 301/614-5772 Home: 410/872-0549
Fax: 301/614-5808 Cellular: 301/613-3782
E-mail: houser@hsb.gsfc.nasa.gov
Web: http://hsb.gsfc.nasa.gov

Dr. Christa Peters-Lidard, Co-PI and Project Manager
Research Scientist, Hydrological Sciences Branch
NASA-GSFC Code 974
Greenbelt, Maryland 20771
Office:301/614-5811 Fax:301/614-5808
E-mail: cpeters@hsb.gsfc.nasa.gov
Web: http://hsb.gsfc.nasa.gov/cpeters

7 List of Publications

None at this time.

8 List of Conference Presentations

Peters-Lidard, C. D., P. R. Houser, M. Rodell, B. Cosgrove, 2003: NASA's Land Data Assimilation Systems, CNES Workshop on Hydrology from Space, 29 Sep-01 Oct 2003, Toulouse, France.

Peters-Lidard, C. D., S. Kumar, Y. Tian, J. L. Eastman, and P. R. Houser, 2004: GLOBAL URBAN-SCALE LAND-ATMOSPHERE MODELING WITH THE LAND INFORMATION SYSTEM, Oral Presentation 4.1 at Symposium on Planning, Nowcasting, and Forecasting in the Urban Zone, American Meteorological Society Annual Meeting, January 2004, Seattle, WA.

9 List of Media References

None at this time.

10 List of Patents Filed

None at this time.

11 List of Students/Post Docs Trained

The LIS project enlisted the help of two undergraduate students this summer to assist with construction of the cluster. In addition, the project hired two post-docs through GEST for the duration of the three-year project (Kumar, Tian), and is also supporting a part-time GEST associate (Eastman).

11.1 Students Trained

- Uttam Majumder, SIECA Summer Intern, 2002
- Nikkia Anderson, Summer Intern, 2002

11.2 Post Docs Trained

- Sujay Kumar, Ph.D.
- Yudong Tian, Ph.D.
- Joseph L. Eastman, Ph.D.

12 References

ALMA: http://www.lmd.jussieu.fr/ALMA/

Atlas, R. M., and R. Lucchesi, File Specification for GEOS-DAS Gridded Output. Available online at : http://dao.gsfc.nasa.gov/DAO_docs/File_Spec_v4.3.html , 2000.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H. Pan, V. Koren, Y. Duan, M. Ek, and A. Betts, "Modeling of land-surface evaporation by four schemes and comparison with FIFE observations", *J. Geophys. Res.*, 101, D3, 7251-7268, 1996.

CLM: http://www.cgd.ucar.edu/tss/clm/

Collatz G. J., C. Grivet, J. T. Ball, and J. A. Berry, J. A. "Physiological and Environmental Regulation of Stomatal Conductance: Photosynthesis and Transpiration: A Model that includes a Laminar Boundary Layer", *Agric. For. Meteorol.* , 5, pp 107 -- 136, 1991.

Derber, J. C., D. F. Parrish, and S. J. Lord, "The new global operational analysis system at the National Meteorological Center", *Wea. And Forecasting,* 6, pp 538-547, 1991.

ESMF: http://www.esmf.ucar.edu/

GrADS-DODS server: http://grads.iges.org/grads/gds/

Hamill, T. M., R. P. d'Entremont, and J. T. Bunting, "A description of the Air Force real-time nephanalysis model", *Wea. Forecasting,* 7, pp 238-306, 1992.

Hofstee, H. P., J. J. Likkien, and J. L. A. Van De Snepscheut "A Distributed Implementation of a Task Pool". *Research Directions in High-Level Parallel Programming Languages*, pp 338--348, 1991.

Jarvis, P. G., " The interpretation of leaf water potential and stomatal conductance found in canopies in the field", *Phil. Trans. R. Soc.* London, Ser. B, 273, pp 593 – 610, 1976.

Kopp, T. J. and R. B. Kiess, "The Air Force Global Weather Central cloud analysis model", *AMS 15th Conf. on Weather Analysis and Forecasting,* Norfolk, VA, pp 220-222, 1996.

LDAS: http://ldas.gsfc.nasa.gov/

NOAH: http://www.emc.ncep.noaa.gov/mmb/gcp/noahlsm/README_2.2.htm

Pfaendtner, J., S. Bloom, D. Lamich, M. Seablom, M. Sienkiewicz, J. Stobbie, and A. da Silva, "Documentation of the Goddard Earth Observing System (GEOS) Data Assimilation System – Version 1", *NASA Technical Memorandum* 104606, 4, pp 44, 1995.

Peters-Lidard, C. D., P. R. Houser, M. Rodell, B. Cosgrove, 2003: NASA's Land Data Assimilation Systems, CNES Workshop on Hydrology from Space, 29 Sep-01 Oct 2003, Toulouse, France.

Peters-Lidard, C. D., S. Kumar, Y. Tian, J. L. Eastman, and P. R. Houser, 2004: GLOBAL URBAN-SCALE LAND-ATMOSPHERE MODELING WITH THE LAND INFORMATION SYSTEM, Oral Presentation 4.1 at Symposium on Planning, Nowcasting, and Forecasting in the Urban Zone, American Meteorological Society Annual Meeting, January 2004, Seattle, WA.

Reynolds, C. A., T. J. Jackson, and W. J. Rawls, "Estimating available water content by linking the FAO Soil Map of the World with global soil profile databases and pedo-transfer functions" *American Geophysical Union, Fall Meeting, Eos Trans. AGU*, 80, 1999.

Richards, L. A., "Capillary conduction of liquids in porous media", *Physics,* 1, pp 318—333, 1931.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, and K. Mitchell, The Global Land Data Assimilation System. Bull. Amer. Meteor. Soc., 2003.

Rogers, E., T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. Baldwin, N. W. Junker, and Y. Lin, "Changes to the operational "early" eta analysis / forecast system at the National Centers for Environmental Prediction" *Wea. Forecasting*, 11, pp 391-413, 1996.

Shapiro, R. "A simple model for the calculation of the flux of direct and diffuse solar radiation through the atmosphere", AFGL-TR-87-0200, Air Force Geophysics Lab, Hanscom AFB, MA.

Turk, F. J., G. Rohaly, J. D. Hawkins, E. A. Smith, A. Grose, F. S. Marzano, A. Mugnai, and V. Levizzani, "Analysis and assimilation of rainfall from blended SSM/I, TRMM, and geostationary satellite data", *AMS 10th Conf. On Sat. Meteor. and Ocean.,* Long Beach, CA, 9-14 January, pp 66-69, 2000.

VIC: http://www.hydro.washington.edu/Lettenmaier/Models/VIC/VIChome.html